

プレースメントテスト用のアダプティブテストを作ろう

## LET'S EXPERIENCE THE ADAPTIVE TEST FOR PLACEMENT

秋山 實, 東北大学大学院教育情報学教育部

Minoru Akiyama, Tohoku University Graduate School Educational Infomatics

akiyama@ei.tohoku.ac.jp

概要：潜在ランク理論に基づくコンピュータ適応型テストは、項目応答理論に基づくものよりも制約が少なく、小規模な利用に向いている。しかし、小規模な利用ではアイテムバンクの特性が偏っていることが多く、システムを最適化しないと期待した結果が得られない場合が多い。このワークショップでは、筆者が開発したプラグインを Moodle に組込んで使用する。参加者は、持参した過去のテストデータを使ってシミュレーションを行い、テストアイテムを再利用してプレースメントテストを行なう一連の体験をハンズオン形式のワークショップで体験できる。参加者にはテストデータとテストアイテムを持参することを推奨する。

キーワード：潜在ランク理論, アダプティブテスト, シミュレーション, Moodle, 項目応答理論

### 潜在ランク理論

古典理論や項目応答理論 (Item Response Theory; IRT) では、受験者の能力を間隔尺度で表す。受験者能力を有効数字 2 ないし 3 桁の数値で表現することが多く、それが能力を高い精度で測定できる、という誤解を生みやすい。荘島 (2007) は、テストの得点 1 点で一喜一憂する風潮を生んでいることを指摘し、テストの精度は受験者の能力をせいぜい 10 レベル程度のランクに順序づけることしかできないという考えのもとに順序尺度を採用する潜在ランク理論 (Latent Rank Theory; LRT) を提案した。LRT では、受験者の能力とテストの問題の特性をランク数に相当する数の要素を持つベクトルで表す。これらの受験者能力と問題の特性パラメータは、ニューラルネットワークの一種である自己組織化マップの性質を利用して推定する。受験者が受験したテストの回答パターンをニューラルネットワークに学習させることにより、テストの回答の正誤のパターンが似ている者同士が能力ランクとしてグルーピングされる。LRT のこの様な特性はプレースメントテストに適しているといえる。プレースメントテストは、その結果によって同じ能力特性の受験者を一つのグループにまとめることが目的であるからだ。

IRT は、受験者の能力をスカラー量として表すのに対し、LRT は、ランク数に応じた要素数を持つベクトルとして表し、各要素はスカラー量であるので情報量は IRT よりも多く、グラフ化すれば試験実施者も受験者も直感的にそれを理解することができる。

一般的にコンピュータ適応型テスト（Computerized Adaptive Test; CAT）は、受験者全員が同じ問題を受験するリニアテストに比較して受験者の能力を回答のたびに推定し、受験者の能力にあった問題を出題するので、受験に要する時間がおおよそ半分以下になることが知られており、これが CAT の第一の特長である。

これまで CAT の基礎として使われる理論は IRT が主流であった。アメリカの多くの公的なテストが IRT を採用し、IRT を基礎とした CAT に関するさまざまなアルゴリズムが提案され、使用されてきた。それらの多くは大規模なテストで十分大きなアイテムバンク（問題のデータベース）を持つものであった。しかし、CAT が普及するに伴い、大学等で独自に CAT を構築するような試みも出てきたが、小規模なアイテムバンクしか用意できない場合、CAT の動作は理論上の動作とは異なり、アイテムバンクを構成する問題の特性に依存した動きとなり、そのアルゴリズムの特徴を活かすことが難しくなってくる。

#### シミュレーションによる最適化

秋山は、LRT に基づく CAT をオープンソースソフトウェアの e ラーニングシステム Moodle のプラグインとして LRT-CAT を開発した。LRT-CAT は、過去のテストの問題の特性データと受験者の能力データおよびその回答データを利用して行うリアルデータシミュレーション機能と乱数を利用して行うモンテカルロシミュレーション機能を持っており、この機能を使ってシミュレーションを行うことで小規模なアイテムバンクであっても、最適な結果を得られる設定（何問で終了させるか、どれくらいの精度を求めるか、どのようなアルゴリズムで問題を選択するか、どのようなアルゴリズムで能力を推定するかなど）を模索し決めることができる。

#### ワークショップの内容

このワークショップでは、ハンズオン形式で実施します。参加者は、主催者側が用意した問題と回答データを Moodle サーバに登録し、リアルデータシミュレーションとモンテカルロシミュレーションを行なって、試行錯誤によって CAT の最適条件を求めるプロセスを体験することができます。冒頭、主催者が LRT-CAT の概要を説明し、選択できるアルゴリズムや条件、操作の方法について概観した上で、参加者は各自 PC を使って Moodle にアクセスし、すべての過程を体験します。

参加者は各自ノート PC を持参することが望ましいが、必ずしも持参していなくてもハンズオンワークショップに参加することができるよう準備する予定です。読みながら操作を進めていけるよう詳細なステップバイステップの操作マニュアルを資料として用意します。Moodle に接続するためのネットワークは、参加者が持参した PC がアクセスできるよう無線 LAN を用意します。